

ICASE

ON THE APPROXIMATION OF EIGENVALUES
ASSOCIATED WITH FUNCTIONAL DIFFERENTIAL EQUATIONS

Kazufumi Ito

(NASA-CR-185810) ON THE APPROXIMATION OF
EIGENVALUES ASSOCIATED WITH FUNCTIONAL
DIFFERENTIAL EQUATIONS (ICASE) 29 p

N89-71365

Unclas
00/64 0224361

Report No. 82-29
September 23, 1982

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE



RESEARCH ASSOCIATION

ON THE APPROXIMATION OF EIGENVALUES
ASSOCIATED WITH FUNCTIONAL DIFFERENTIAL EQUATIONS

Kazufumi Ito
Institute for Computer Applications in Science and Engineering

Abstract

In this paper an approximation method based upon spline functions is developed for the eigenvalue problem associated with functional differential equations. Convergence results are established and the rate of convergence is investigated. Numerical results for cubic and quintic spline based methods are given. The paper concludes with a brief discussion of other possible approximation methods.

Research for this work was supported under NASA Contract No. NAS1-15810 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

Introduction

In recent years spectral approximation methods for elliptic differential operators have been extensively treated (see [12], [14] and [16] for a survey of the literature).

In this paper spectral approximation results are established for a large class of linear functional differential equations (FDEs). The eigenvalue problem for FDEs is formulated as the eigenvalue problem of a linear closed operator A in the product space. We consider the least squares-type approximation for $T = A^{-1}$ via spline approximation technique and the general results on spectral approximation for compact operator [1], [16] are then applied in our situation.

A more direct and natural method would involve the approximation of A as in Banks and Kappel [4], and as discussed below we have carried out computations based on such a scheme. However, our scheme performs much better numerically. Moreover we can establish convergence results for our scheme, whereas we have been unable to establish convergence of the methods in the case of the more direct approximation of A .

The knowledge of the spectrum plays an important role in studying the stability and controllability, etc., of evolution systems [13] and gives an approximation scheme for such systems [5].

In the last several years it has been widely recognized that the product space ($Z = \mathbb{R}^n \times L_2$) provides an appropriate state space for the investigation of certain problems involving FDEs of the retarded type as well as of the neutral type [6], [9], [11]. These spaces have been used in the context of optimal control and estimation problems for systems governed by FDEs (see [2], [9], [10], [13], for a survey of literature), and approximation techniques of identification and optimal control for retarded systems, e.g., [2], [3], [4], [8].

We first summarize in Section 2 certain theoretical results about FDEs in order to set up our eigenvalue approximation problem in $\mathbb{R}^n \times L_2$ and give a complete discussion of our ideas. A general approximation of the resolvent of A is developed in Section 3. In Section 4 we study a realization of the general scheme given in Section 3 by choosing spaces of spline functions as approximate subspaces. Finally in Section 5, numerical results for cubic and quintic spline based methods are presented. For comparison, we also give a result for the spline approximation of A which is discussed in [4].

Throughout the paper, the dimension n and the delay $r \in (0, \infty)$ are assumed to be fixed and the following notation will be used. The Hilbert space of \mathbb{R}^n -valued, square integrable functions on the interval $[-r, 0]$ is denoted by L_2 . C^k denotes the space of \mathbb{R}^n -valued continuous functions which possess k continuous derivatives on $[-r, 0]$. For $k = 0$ this is usual space of continuous functions which is denoted simply by C . W_2^k is the Sobolev space, the space of \mathbb{R}^n -valued functions f on $[-r, 0]$ with $f^{(k-1)}$ absolutely continuous and $f^{(k)} \in L_2$. We denote by Z the product space $\mathbb{R}^n \times L_2$. Given an element $\phi \in Z$, $\phi^0 \in \mathbb{R}^n$ and $\phi^1 \in L_2$ denote the two coordinates of ϕ , $\phi = (\phi^0, \phi^1)$. The symbol $\langle\langle \cdot, \cdot \rangle\rangle$ stands for natural inner product in Z . All other inner products are denoted by $\langle \cdot, \cdot \rangle$ when the underlying space can be understood from the context. $\|\cdot\|$ denotes the norm of elements of a Banach space and of operators between Banach spaces and $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^n . For a linear operator A we use the standard notation: $\mathcal{D}(A)$, $R(A)$, $N(A)$, for the domain, range and null space of A , respectively, and $\rho(A)$ and $\sigma(A)$ denote the resolvent set and spectrum of A , respectively. Given a measurable function $x : [-r, 0] \rightarrow \mathbb{R}^n$ and $t > 0$, the function $x_t : [-r, 0] \rightarrow \mathbb{R}^n$ is defined by $x_t(\theta) = x(t+\theta)$, $\theta \in [-r, 0]$.

2. The Linear Functional Differential Equation

In this section we state the type of equation we consider and the results for this equation which are important for the discussions to follow.

Given $(\eta, \phi) \in Z$ we consider the functional integral equation in \mathbb{R}^n

$$(2.1) \quad Dx_t = \eta + \int_0^t Lx_s ds$$

$$x_0 = \phi$$

where D and L are linear \mathbb{R}^n -valued operators with domains $\mathcal{D}(D)$ and $\mathcal{D}(L)$ both subspaces of L_2 . Assume that D and L are given by

$$(2.2) \quad D\phi = \phi(0) - \sum_{i=1}^m A_{-i} \phi(\theta_i) - \int_{-r}^0 A^1(\theta) \phi(\theta) d\theta$$

and

$$(2.3) \quad L\phi = \sum_{i=0}^m A_i \phi(\theta_i) + \int_{-r}^0 A^2(\theta) \phi(\theta) d\theta$$

where $-r = \theta_m < \dots < \theta_0 = 0$ and $A_i, A^1(\theta)$ are $n \times n$ matrices, the element of $\theta \rightarrow A^1(\theta)$ being integrable on $[-r, 0]$ for $i = 1, 2$.

For $(\eta, \phi) \in Z$ there exists a unique solution x to (2.1) in $L_2([-r, \tau], \mathbb{R}^n)$ for each $\tau > 0$ such that $(Dx_t, x_t) \in Z$ for each $t \in [0, \tau]$ and depends continuously on the initial data (η, ϕ) (e.g. see [6], [9]). Define for $t > 0$, $S(t) : Z \rightarrow Z$ by $S(t)(\eta, \phi) = (Dx_t, x_t)$ where x is the solution to (2.1). Then $\{S(t), t > 0\}$ forms a strongly continuous semigroup in Z . Using rather standard arguments (e.g., [6], [9], and [18]) one can then show:

Theorem 2.1

(i) If A denotes the infinitesimal generator of $S(t)$, $t > 0$, then

$$(2.4) \quad \mathcal{D}(A) = \{(\eta, \phi) \in Z \mid \dot{\phi} \in L_2 \text{ and } \eta = D\phi\},$$

and for $(D\phi, \phi) \in \mathcal{D}(A)$,

$$(2.5) \quad A(D\phi, \phi) = (L\phi, \dot{\phi}).$$

(ii) The spectrum of A is all point spectrum and $\lambda \in \sigma(A)$ if $\det \Delta(\lambda) = 0$ where

$$(2.6) \quad \Delta(\lambda) = \lambda D(e^{\lambda \cdot}) + L(e^{\lambda \cdot}).$$

For each $\lambda \in \rho(A)$, the resolvent of A is compact and is given by

$$(\lambda I - A)^{-1} z = (D\psi, \psi) \quad \text{for } z = (\eta, \phi) \in Z,$$

with

$$(2.7) \quad \psi(\theta) = e^{\lambda \theta} b + \int_{\theta}^0 e^{\lambda(\theta-\xi)} \phi(\xi) d\xi,$$

where

$$(2.8) \quad b = \Delta^{-1}(\lambda) \left[\eta + (\lambda D + L) \left(\int_0^{\cdot} e^{\lambda(\cdot-\xi)} \phi(\xi) d\xi \right) \right].$$

(iii) If $\lambda \in \sigma(A)$ is a zero of $\det \Delta(\lambda)$ of multiplicity m , then $Z = N(\lambda I - A)^m + R(\lambda I - A)^m$ and both $N(\lambda I - A)^m$ and $R(\lambda I - A)^m$ are

invariant with respect to $S(t)$, $t \geq 0$. Moreover $N(\lambda I - A)^m$ has dimension m and its element is represented in the form $(D\phi, \phi)$ with $\phi = \left(\sum_{j=1}^m \alpha_j \theta^j \right) e^{\lambda \theta}$, $\alpha_j \in \mathbb{R}^n$.

Remark 2.2 If $(\eta, \phi) \in \mathcal{D}(A)$ then the solution x to (2.1) is absolutely continuous in $[-r, \tau]$ and satisfies the differential version of equation (2.1):

$$(2.9) \quad \frac{d}{dt} D x_t = L x_t \quad x_0 = \phi$$

for almost all t .

Remark 2.3 Let us introduce an equivalent norm in Z such that Z with this norm is again a Hilbert space, i.e., define the weighting function g to be step function on $[-r, 0]$ such that

$$g(\theta) = j \quad \text{for } \theta \in (\theta_{m-j+1}, \theta_{m-j}) \quad j = 1, \dots, m,$$

and let us denote by Z_g the completion of Z with respect to inner product

$$\langle (\eta_1, \phi_1), (\eta_2, \phi_2) \rangle_{Z_g} = \langle \eta_1, \eta_2 \rangle_{\mathbb{R}^n} + \int_{-r}^0 \langle \phi_1(\theta), \phi_2(\theta) \rangle_{\mathbb{R}^n} g(\theta) d\theta.$$

Then for retarded systems (i.e., $D\phi = \phi(0)$ in (2.2)), it is shown in [4] that $A - \beta I$ is dissipative in Z_g for some positive constant β . Obviously, the norm $\|\cdot\|_{Z_g}$ is equivalent to $\|\cdot\|_Z$. Therefore, all results stated in Theorem 2.1 remain valid in Z_g .

3. An Approximation of the Resolvent

Without loss of generality we assume that $T = A^{-1}$ exists, i.e., $0 \in \rho(A)$ since adding a constant c only shifts the eigenvalues of A . It is easily shown that $\lambda \in \sigma(A)$ if $\mu = 1/\lambda \in \sigma(T)$ and moreover, the principal vectors of A associated with λ are the same as the principal vectors of T associated with μ . Under this assumption there exists a positive constant ω such that

$$(3.1) \quad \langle\langle Az, Az \rangle\rangle > \omega \|z\|^2 \quad \text{for all } z \in \mathcal{D}(A).$$

Indeed, for any element $z \in \mathcal{D}(A)$ there exists a unique $\phi \in Z$ such that $z = A^{-1}\phi$ and therefore

$$\langle\langle Az, Az \rangle\rangle = \langle\langle \phi, \phi \rangle\rangle = \|\phi\|^2.$$

However, there exists a positive constant α such that

$$(3.2) \quad \|A^{-1}\phi\| \leq \alpha \|\phi\| \quad \text{for all } \phi \in Z,$$

so it follows that $\langle\langle Az, Az \rangle\rangle \geq (1/\alpha^2) \|z\|^2$.

Consider the following approximation of T : if $\{Z_N\}$ is a sequence of subspaces of Z , then find a minimizing element in Z_N of

$$(3.3) \quad \|Az - f\| \quad \text{for each } f \in Z.$$

Then such an element z^N must satisfy

$$(3.4) \quad \langle\langle Az^N - f, A\phi \rangle\rangle = 0 \quad \text{for all } \phi \in Z^N.$$

Hence from (3.1) it follows that the minimization problem (3.3) has a unique solution z^N satisfying (3.4). Let us define a linear operator T^N in Z by

$$(3.5) \quad T^N f = z^N \quad \text{for } f \in Z$$

where z^N is the unique solution of (3.4).

Lemma 3.1

Suppose the following conditions are satisfied.

- (a) $Z^N \subset \mathcal{D}(A)$
- (b) There exists a sequence $\{\phi_N\}$ such that $\phi_N = (D\phi_N, \phi_N) \in Z^N$,
 $\lim L \phi_N = L\phi$ in R^n and $\lim \dot{\phi}_N = \dot{\phi}$ in L_2 for all $\phi \in C^k$, $k > 2$.

Then

- (i) $T^N \rightarrow T$ strongly in Z .
- (ii) $\{T^N\}$ is collectively compact, i.e., the set $\{T^N f : \|f\| < 1, N = 1, 2, \dots\}$ is sequentially compact.

Proof:

(i) Let $f = (f^0, f^1) \in Z$. Since C^{k-1} is dense in L_2 for any $\epsilon > 0$ there exists a $\hat{f}^1 \in C^{k-1}$ such that $\|f^1 - \hat{f}^1\| < \epsilon$. Let us define $\hat{f} = (f^0, \hat{f}^1) \in Z$ and $\hat{\phi} = A^{-1} \hat{f} = (D\hat{\phi}, \hat{\phi})$. It follows from (2.7) and (2.8) that $\hat{\phi} \in C^k$. It then follows from the condition (b) there exists an N_0

and a sequence $\{\phi_N\}$ such that if $N > N_0$ then

$$\|A(\phi_N - \hat{\phi})\|_2 < \varepsilon.$$

Note that

$$\|A(T^N f - Tf)\| < \|A(\phi_N - \hat{\phi}) + \hat{f} - f\| < 2\varepsilon.$$

Hence from (3.1) we have $\|T^N f - Tf\| < \frac{2\varepsilon}{\omega}$.

(ii) For any $f \in Z$

$$\|AT^N f\| < \|(AT^N f - f) + f\| < \|AT^N f - f\| + \|f\|.$$

By the definition of T^N

$$< \|f\| + \|f\| = 2\|f\|.$$

Hence $\{AT^N f : \|f\| \leq 1, N = 1, 2, \dots\}$ is a bounded set in Z . Therefore (ii) follows from the compactness of A^{-1} . (Q. E. D.)

Let μ be a fixed nonzero eigenvalue of T with algebraic multiplicity m and suppose that Γ is circle centered at μ which lies in $\rho(T)$ and which encloses no other points of $\sigma(T)$. Then the operator

$$(3.6) \quad E = \frac{1}{2\pi i} \int_{\Gamma} (zI - T)^{-1} dz$$

is a projection operator onto $R(E) = N((\mu I - T)^m)$. The following results for the case of strong convergence of collectively compact operators can be found

in Anselone [1]. For N sufficiently large $\Gamma \subset \rho(T^N)$ and the spectral projection

$$(3.7) \quad E^N = \frac{1}{2\pi i} \int_{\Gamma} (zI - T^N)^{-1} dz$$

exists; E^N converges strongly to E and $\{E^N\}$ is collectively compact; and $\dim R(E^N) = \dim R(E) = m$. Thus, counting according to algebraic multiplicities, there are m eigenvalues of T^N in Γ which we denote $\mu_1(N), \dots, \mu_m(N)$. For each j , $\lim \mu_j(N) = \mu$ as $N \rightarrow \infty$.

Moreover, the general results of collectively compact convergence of a family of compact operators in Osborn [16] provide the following estimates for the convergence. Given two closed subspaces P and Q of Z we define

$$(3.8) \quad \delta(P, Q) = \sup_{\substack{z \in P \\ \|z\| = 1}} \text{dist}(z, Q).$$

Note that $\delta(P, Q) = \delta(Q, P)$ since Z is a Hilbert space.

Theorem 3.2

(i). There exists a constant C_1 such that

$$\delta(R(E), R(E^N)) \leq C_1 \| (T - T^N)|_{R(E)} \|$$

for all sufficiently large N , where $(T - T^N)|_{R(E)}$ denotes the restriction of $T - T^N$ to $R(E)$.

(ii). Define the arithmetic mean of $\{\mu_j(N)\}$:

$$\hat{\mu}(N) = \frac{1}{m} \sum_{j=1}^m \mu_j(N).$$

Then there exists a constant C_2 such that

$$|\mu - \hat{\mu}(N)| \leq C_2 \|(T - T^N)|_{R(E)}\|.$$

(iii). Let α be the ascent of $\mu I - T$. Then there exists a constant C_3 such that

$$|\mu - \mu_j(N)|^\alpha \leq C_3 \|(T - T^N)|_{R(E)}\|.$$

(iv). Let $\mu(N)$ be an eigenvalue of T^N and w^N a unit vector in $R(E^N)$ satisfying $(\mu(N) - T^N)^k w^N = 0$ for some positive integer $k < \alpha$. Then for any integer ℓ with $k < \ell < \alpha$, there exists a vector $u^N \in R(E)$ such that $(\mu I - T)^\ell u^N = 0$ and

$$\|u^N - w^N\| \leq C_4 \|(T - T^N)|_{R(E)}\|^{(\ell-k+1)/\alpha}.$$

4. Spline Approximation

In this section we choose the Z^N in Section 3 as a certain subspace of spline functions and we discuss the rate of convergence of the approximation scheme.

Let $\{e_j^N\}$, $j = 1, \dots, N + 2k + 1$ be the scalar $(2k - 1)$ th order spline function on $[-r, 0]$ corresponding to the partition $t_j^N = -j(r/N)$, $j = 0, \dots, N$ of $[-r, 0]$ and Z_{2k-1}^N be the linear subspace of Z spanned by elements of the form $(D\phi, \phi) \in Z$ with

$$\phi = (0, \dots, 0, e_j^N, 0, \dots, 0).$$

Then $z_{2k-1}^N \subset \mathcal{D}(A)$ and $\dim z_{2k-1}^N = n(N + 2k + 1) = k_N$.

Lemma 4.1

$\{z_{2k-1}^N\}$ satisfies the conditions of Lemma 3.1 for $k > 1$ and there exists a constant C_5 depending upon E such that

$$(4.1) \quad \|(T - T^N)|_{\mathcal{R}(E)}\| \leq C_5 \left(\frac{1}{N}\right)^{2k-1}.$$

Proof: Let $z = (D\phi, \phi) \in \mathcal{D}(A)$ with $\phi \in C^{2k}$ and ϕ_I^N denote the interpolating $(2k-1)$ th order spline function defined by

$$(4.2) \quad \phi_I^N(t_j^N) = \phi(t_j^N) \quad j = 0, \dots, N$$

$$\phi_I^{N(1)}(0) = \phi^{(1)}(0)$$

and

$$\phi_I^{N(1)}(-r) = \phi^{(1)}(-r), \quad i = 1, \dots, k-1.$$

Using the well-known convergence properties of interpolating spline, e.g., [17, p. 3]

$$(4.3) \quad \|\dot{\phi}_I^N - \dot{\phi}\| \leq M \left(\frac{1}{N}\right)^{2k-1} \|\phi\|_{W_2^{2k}}$$

for some positive constant M .

For $\theta \in [-r, 0]$ we have

$$\phi_I^N(\theta) = \phi(0) + \int_0^\theta \dot{\phi}_I^N(\xi) d\xi$$

and hence

$$\begin{aligned}
 |\phi_I^N(\theta) - \phi(\theta)| &\leq \int_{\theta}^0 |\dot{\phi}_I^N(\xi) - \dot{\phi}(\xi)| d\xi \\
 &\leq r^{1/2} \|\dot{\phi}_I^N - \dot{\phi}\|_{L_2} \\
 &\leq r^{1/2} M \left(\frac{1}{N}\right)^{2k-1} \|\phi\|_{W_2^{2k}}.
 \end{aligned}$$

Note that L is a continuous functional $C \rightarrow \mathbb{R}^N$. Thus, the condition (b) of Lemma 3.1 is satisfied and moreover, we have

$$\|A(z - z_I^N)\| \leq M(1 + r^{1/2} \|L\|) \left(\frac{1}{N}\right)^{2k-1} \|\phi\|_{W_2^{2k}},$$

where $z_I^N = (D\phi_I^N, \phi_I^N) \in Z$ and

$$(4.4) \quad \|L\| = \sum_{i=0}^m |A_i| + \int_{-r}^0 |A^1(\theta)| d\theta.$$

For any $f = (f^0, f^1) \in R(E)$ for a fixed μ let $z = Tf = (D\phi, \phi)$, then we have

$$\begin{aligned}
 \|(T - T^N)f\| &\leq \frac{1}{\omega} \|A(T - T^N)f\| \\
 &\leq \frac{1}{\omega} \|A(z - z_I^N)\| \\
 &\leq \frac{M}{\omega} (1 + r^{1/2} \|L\|) \left(\frac{1}{N}\right)^{2k-1} \|\phi\|_{W_2^{2k}} \\
 &\leq \frac{M}{\omega} (1 + r^{1/2} \|L\|) \left(\frac{1}{N}\right)^{2k-1} \|f^1\|_{W_2^{2k-1}}.
 \end{aligned}$$

From (iii) of Theorem 2.1 it follows that $\|f^1\|_{W_2^{2k-1}} \leq \gamma \|f^1\|_{L_2}$ for some

positive constant γ depending upon f^1 . Therefore (4.1) follows from the finiteness of $\dim R(E)$. (Q. E. D.)

From the estimate (4.1) and Theorem 3.2 the rate of convergence for eigenvalues is at least $O((\frac{1}{N})^{2k-1})$.

According to the equation (3.4), $T^N|_{Z_{2k-1}^N}$ has the following coordinate representation. Let β^N be given by

$$(4.5) \quad \beta^N = (e_1^N, \dots, e_{N+2k+1}^N) \otimes I,$$

when \otimes denotes the Kronecker product and I is the $n \times n$ identity matrix and

$$(4.6) \quad \hat{\beta}^N = (D\beta^N, \beta^N).$$

Define $k_N \times k_N$ matrices

$$(4.7) \quad R_{2k-1}^N = \langle\langle A\hat{\beta}^N, A\hat{\beta}^N \rangle\rangle,$$

and

$$(4.8) \quad H_{2k-1}^N = \langle\langle \hat{\beta}^N, A\hat{\beta}^N \rangle\rangle.$$

Then

$$(4.9) \quad T^N(\hat{\beta}^N \alpha) = \hat{\beta}^N (R_{2k-1}^N)^{-1} (H_{2k-1}^N)^T \alpha,$$

where α is k_N dimensional column vector.

Note that from (3.1) it is easy to show that R_{2k-1}^N is invertible and that $\sigma(T^N)$ is equal to the set of eigenvalues of the matrix:

$$(4.10) \quad X_{2k-1}^N = (R_{2k-1}^N)^{-1} (H_{2k-1}^N)^T.$$

5. Numerical Results and Remarks

In this section we present some numerical results to illustrate the approximation methods described in Section 3 and 4 and discuss other possible approximation methods.

The matrices (4.7), (4.8), the inverse of R_{2k-1}^N and the eigenvalues of the matrix X_{2k-1}^N were computed using standard algorithms such as Gaussian quadrature, Gauss elimination and the QR-method [7].

In order to analyze the accuracy of approximation we need the approximate reference eigenvalues. To this end the approximate eigenvalues obtained by our methods were put into an IMSL package, which computes the roots of an analytic complex function using Muller's method with deflation, as initial guesses. In the tables below, λ_{EXACT}^N denotes the eigenvalues obtained by the following steps.

- Step 1: Compute the approximate eigenvalues by using a quintic spline based method with N appropriately chosen for each example.
- Step 2: Compute the roots of the characteristic equation by using Muller's method in which the eigenvalues obtained in Step 1 are used as initial guesses. A root is accepted if two successive approximations to it agree in the first seven digits.

Example 1

For the first example we consider the first order equation:

$$\dot{x}(t) = x(t) + x(t-1).$$

For this example

$$\mu = \frac{1}{\lambda} \in \sigma(T) \quad \text{if} \quad \Delta(\lambda) = \lambda - 1 - e^{-\lambda} = 0$$

and the calculations were carried out for cubic and quintic spline methods. In the tables below λ_c^N and λ_q^N denote the computed eigenvalues based on cubic and quintic splines respectively. δ_c^N and δ_q^N are the errors $|\Delta(\lambda_c^N)|^2$ and $|\Delta(\lambda_q^N)|^2$ respectively. The results clearly show that the convergence of λ_q is much faster than the convergence of λ_c . For both methods the rate of convergence is better than we expected in Section 4 and the eigenvalues with small modulus are obtained almost exactly regardless of N .

TABLE I

No.	λ_c^8	δ_c^8	λ_c^{16}	δ_c^{16}
1	1.278465	3×10^{-20}	1.278465	5×10^{-24}
2	$-1.588318 \pm 4.155300 \text{ i}$	5×10^{-10}	$-1.588317 \pm 4.155305 \text{ i}$	1×10^{-13}
3	$-2.422831 \pm 10.68360 \text{ i}$	4×10^{-3}	$-2.417678 \pm 10.68602 \text{ i}$	3×10^{-7}
4	$-3.247468 \pm 17.03489 \text{ i}$	7×10^1	$-2.863190 \pm 17.05637 \text{ i}$	9×10^{-4}
5			$-3.190732 \pm 23.39339 \text{ i}$	3×10^{-1}
6			$-3.599035 \pm 29.79009 \text{ i}$	5×10^1

No.	λ_c^{32}	δ_c^{32}	$\lambda_{\text{EXACT}}^{94}$
1	1.278465	3×10^{-24}	1.278465
2	$-1.588317 \pm 4.155305 \text{ i}$	3×10^{-17}	$-1.588317 \pm 4.155305 \text{ i}$
3	$-2.417631 \pm 10.68603 \text{ i}$	5×10^{-11}	$-2.417631 \pm 10.68603 \text{ i}$
4	$-2.861518 \pm 17.05612 \text{ i}$	9×10^{-8}	$-2.861502 \pm 17.05611 \text{ i}$
5	$-3.167905 \pm 23.38566 \text{ i}$	2×10^{-5}	$-3.167754 \pm 23.38558 \text{ i}$
6	$-3.402828 \pm 29.69862 \text{ i}$	1×10^{-3}	$-3.401945 \pm 29.69798 \text{ i}$
7	$-3.595517 \pm 36.00485 \text{ i}$	4×10^{-2}	$-3.591627 \pm 36.00146 \text{ i}$
8	$-3.765356 \pm 42.31378 \text{ i}$	7×10^{-1}	$-3.751047 \pm 42.29965 \text{ i}$
9	$3.934929 \pm 48.64474 \text{ i}$	1×10^1	$-3.888543 \pm 48.59442 \text{ i}$

TABLE II

No.	λ_q^8	δ_q^8	λ_q^{16}	δ_q^{16}
1	1.278465	4×10^{-25}	1.278465	7×10^{-25}
2	$-1.588317 \pm 4.155305 \text{ i}$	3×10^{-18}	$-1.588317 \pm 4.155305 \text{ i}$	3×10^{-22}
3	$-2.417631 \pm 10.68600 \text{ i}$	1×10^{-7}	$-2.417631 \pm 10.68603 \text{ i}$	1×10^{-14}
4	$-2.877073 \pm 17.03424 \text{ i}$	2×10^{-1}	$-2.861504 \pm 17.05611 \text{ i}$	2×10^{-9}
5			$-3.167935 \pm 23.38550 \text{ i}$	2×10^{-5}
6			$-3.408216 \pm 29.69662 \text{ i}$	3×10^{-2}
7			$-3.726049 \pm 35.99129 \text{ i}$	3×10^1

No.	λ_q^{32}	δ_q^{32}	$\lambda_{\text{EXACT}}^{94}$
1	1.278465	4×10^{-24}	1.278465
2	$-1.588317 \pm 4.155305 \text{ i}$	8×10^{-22}	$-1.588317 \pm 4.155305 \text{ i}$
3	$-2.417631 \pm 10.68603 \text{ i}$	1×10^{-20}	$-2.417631 \pm 10.68603 \text{ i}$
4	$-2.861502 \pm 17.05611 \text{ i}$	4×10^{-16}	$-2.861502 \pm 17.05611 \text{ i}$
5	$-3.167754 \pm 23.38558 \text{ i}$	1×10^{-12}	$-3.167754 \pm 23.38558 \text{ i}$
6	$-3.401946 \pm 29.69798 \text{ i}$	7×10^{-10}	$-3.401945 \pm 29.69798 \text{ i}$
7	$-3.591638 \pm 36.00146 \text{ i}$	2×10^{-7}	$-3.591627 \pm 36.00146 \text{ i}$
8	$-3.751138 \pm 42.29968 \text{ i}$	2×10^{-5}	$-3.751047 \pm 42.29965 \text{ i}$
9	$-3.889167 \pm 48.59469 \text{ i}$	1×10^{-3}	$-3.888543 \pm 48.59442 \text{ i}$
10	$-4.013041 \pm 54.88863 \text{ i}$	5×10^{-2}	$-4.009422 \pm 54.88686 \text{ i}$
11	$-4.135839 \pm 61.18780 \text{ i}$	2×10^0	$-4.117267 \pm 61.17761 \text{ i}$
12	$-4.301116 \pm 67.51884 \text{ i}$	5×10^1	$-4.214618 \pm 67.46710 \text{ i}$

According to the calculations in Step 1 with $N = 94$, fifty-three "good" approximations, which agree at least in the first three digits of the real part and four digits of the imaginary part with the reference eigenvalues $\lambda_{\text{EXACT}}^{94}$ were obtained. In addition to these eigenvalues, twenty-two more eigenvalues were obtained through Step 2.

Example 2

For the next example we consider the first order neutral system

$$\dot{x}(t) - \frac{1}{2} \dot{x}(t-1) = x(t) + x(t-1)$$

with the characteristic equation $\det \Delta(x) = \lambda(1 - \frac{1}{2}e^{-\lambda}) - (1 + e^{-\lambda})$. The numerical results for this example were carried out using quintic spline approximation with $N = 94$ and can be found in Table III. We can observe the neutral chain $\{\lambda \in \mathbb{C} \mid \lambda = -\log 2 + i2k\pi, k \in \mathbb{N}\}$. Although we have not listed all of the computed eigenvalues, a total of eighty "good" approximations were obtained via Step 1 and then Step 2 with $N = 94$.

TABLE III

No.	λ_q^{94}	δ_q^{94}	δ_q^{94}	$\lambda_{\text{EXACT}}^{94}$
1	1.414842	3×10^{-22}	1.414842	
2	-2.553133	1×10^{-19}	-2.553133	
3	-0.7108952 \pm 5.775593 i	6×10^{-21}	-0.7108952 \pm 5.775593 i	
4	-0.6969801 \pm 12.32420 i	8×10^{-20}	-0.6969801 \pm 12.32420 i	
5	-0.6948095 \pm 18.68940 i	5×10^{-19}	-0.6948095 \pm 18.68940 i	
6	-0.6940744 \pm 25.01295 i	4×10^{-18}	-0.6940744 \pm 25.01295 i	
7	-0.6937383 \pm 31.32022 i	3×10^{-17}	-0.6937383 \pm 31.32022 i	
\vdots				
19	-0.6932011 \pm 106.7861 i	2×10^{-5}	-0.6931980 \pm 106.7861 i	
20	-0.6931989 \pm 113.0709 i	7×10^{-5}	-0.6931925 \pm 113.0708 i	
21	-0.6932007 \pm 119.3555 i	4×10^{-4}	-0.6931879 \pm 119.3554 i	
22	-0.6932091 \pm 125.6402 i	2×10^{-3}	-0.6931839 \pm 125.6398 i	
23	-0.6932286 \pm 131.9248 i	7×10^{-3}	-0.6931801 \pm 131.9242 i	
24	-0.6932678 \pm 138.2096 i	3×10^{-2}	-0.6931775 \pm 138.2084 i	
25	-0.6933413 \pm 144.4949 i	1×10^{-1}	-0.6931749 \pm 144.4925 i	
26	-0.6934740 \pm 150.7810 i	4×10^{-1}	-0.6931725 \pm 150.7766 i	
27	-0.6937085 \pm 157.0687 i	2×10^0	-0.6931707 \pm 157.0605 i	
28	-0.6941158 \pm 163.3592 i	5×10^0	-0.6931689 \pm 163.3445 i	
29	-0.6948143 \pm 169.6545 i	2×10^1	-0.6931673 \pm 169.6283 i	
30	-0.6959983 \pm 175.9581 i	7×10^1	-0.6931659 \pm 175.9121 i	
31	-0.6979865 \pm 182.2758 i	2×10^2	-0.6931646 \pm 182.1959 i	

Example 3

Here we consider the equation for an oscillator with delayed damping

$$\ddot{x}(t) + \dot{x}(t - 1) + x(t) = 0.$$

Rewriting the above equation as a first order system we have

$$\begin{bmatrix} \dot{x}_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix},$$

where $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. For this example the characteristic equation is given by $\lambda^2 + \lambda e^{-\lambda} + 1 = 0$ and the calculations were carried out using a quintic spline approximation with $N = 45$ which provided forty-one accurate approximation for the eigenvalues through the two-step procedure outlined above. Note that this system has two unstable eigenvalues "0.0219316 \pm 1.601953 i." But the system: $\ddot{x}(t) + \dot{x}(t) + x(t) = 0$ is stable.

TABLE IV

No.	λ_q^{45}	δ_q^{45}	$\lambda_{\text{EXACT}}^{45}$
1	-0.7384324	8×10^{-21}	-0.7384324
2	$0.0219316 \pm 1.601953 \text{ i}$	2×10^{-22}	$0.0219316 \pm 1.601953 \text{ i}$
3	$-2.046879 \pm 7.582028 \text{ i}$	3×10^{-19}	$-2.046879 \pm 7.582028 \text{ i}$
4	$-2.648399 \pm 13.94769 \text{ i}$	3×10^{-17}	$-2.648399 \pm 13.94769 \text{ i}$
5	$-3.017915 \pm 20.27187 \text{ i}$	9×10^{-15}	$-3.017915 \pm 20.27187 \text{ i}$
6	$-3.286399 \pm 26.58018 \text{ i}$	1×10^{-11}	$-3.286399 \pm 26.58018 \text{ i}$
7	$-3.497613 \pm 32.88056 \text{ i}$	5×10^{-9}	$-3.497613 \pm 32.88056 \text{ i}$
8	$-3.671812 \pm 39.17634 \text{ i}$	8×10^{-7}	$-3.671811 \pm 39.17634 \text{ i}$
9	$-3.820082 \pm 45.46920 \text{ i}$	6×10^{-5}	$-3.820078 \pm 45.46920 \text{ i}$
10	$-3.949172 \pm 51.76008 \text{ i}$	3×10^{-3}	$-3.949154 \pm 51.76007 \text{ i}$
11	$-4.063528 \pm 58.04959 \text{ i}$	1×10^{-1}	$-4.063448 \pm 58.04954 \text{ i}$
12	$-4.166328 \pm 64.33820 \text{ i}$	3×10^0	$-4.166003 \pm 64.33796 \text{ i}$
13	$-4.260209 \pm 70.62655 \text{ i}$	6×10^1	$-4.259009 \pm 70.62558 \text{ i}$
14	$-4.348215 \pm 76.91612 \text{ i}$	1×10^3	$-4.344095 \pm 76.91258 \text{ i}$
15	$-4.435822 \pm 83.21115 \text{ i}$	2×10^4	$-4.422504 \pm 83.19908 \text{ i}$
16	$-4.536140 \pm 89.52395 \text{ i}$	2×10^5	$-4.495209 \pm 89.48519 \text{ i}$

Remark 5.1

The approximate methods developed in Banks-Kappel [4] provide a method to approximate the eigenvalues of FDEs of retarded type. Let $\{P^N\}$ be the sequence of orthogonal projections $P^N : Z_g \rightarrow Z_{2k-1}^N$ for a fixed k (see Remark 2.4 for the definition of Z_g) and $A^N = P^N A P^N$. Then it is known that for λ

sufficiently large $(\lambda I - A^N)^{-1}$ converges strongly to $(\lambda I - A)^{-1}$. From this, when combined with the fact that $(\lambda I - A)^{-1}$ is compact, one can argue the convergence results as in Theorem 3.2. The author has tried repeatedly but unsuccessfully to prove that $\{(\lambda I - A^N)^{-1}\}$ is collectively compact. The following numerical results suggest that the convergence results are true. The calculations were carried out for Example 1 by using a quintic spline approximation with $N = 32$. Although we have not listed the results of this scheme for other examples, we note that the convergence of this scheme is much slower than that for our scheme. For instance, the equation errors δ_{B-K}^{32} are much bigger than δ_q^8 and δ_q^{16} in Table II for the eigenvalues with small modulus in this calculations.

TABLE V

No.	λ_{B-K}^{32}	δ_{B-K}^{32}	λ_{EXACT}^{94}
1	1.278464	8×10^{-15}	1.278465
2	-1.588319 \pm 4.155304 i	1×10^{-10}	-1.588317 \pm 4.155305 i
3	-2.417643 \pm 10.68603 i	2×10^{-8}	-2.417631 \pm 10.68603 i
4	-2.861526 \pm 17.05611 i	1×10^{-7}	-2.861502 \pm 17.05611 i
5	-3.167977 \pm 23.38543 i	4×10^{-5}	-3.167754 \pm 23.38558 i
6	-3.403925 \pm 29.69778 i	4×10^{-3}	-3.401945 \pm 29.69798 i
7	-3.600767 \pm 36.00608 i	1×10^{-1}	-3.591627 \pm 36.00146 i
8	-3.772276 \pm 42.33761 i	4×10^0	-3.751047 \pm 42.29965 i
9	-3.861764 \pm 48.75268 i	6×10^1	-3.888543 \pm 48.59442 i

Remark 5.2

From the numerical results above our method appears to yield good approximations for the eigenvalues associated with FDEs. Even for small N (e.g. $N = 8$) it yields good approximations to eigenvalues with small modulus. And it provides good initial guesses for a subroutine to compute the roots of the characteristic equation. In the case when n , the dimension of the system (2.1) is large we have to solve the eigenvalue problem for large order matrix. For such a case it seems better to choose the higher order spline elements for the approximation rather than increasing N . Note that Theorem 3.2 remains valid in Z_g for any weighting function g . For the case when (2.1) has more than one point delay the choice of g as in (2.10) would provide good approximation methods.

For the case when (2.1) only has single point delay, the implementation of the algorithms for our scheme is almost as easy for the scheme described in Remark 5.1. For general ease the inversion of the matrix R_{2k-1}^N is more complicated, so we would use the QZ-method [15] to find eigenvalues of generalized eigenvalue problem: $\lambda R_{2k-1}^N \alpha = H_{2k-1}^N \alpha$.

Let us discuss other possible approximation methods.

Remark 5.3

Let us consider a sequence of approximations to T defined by $P^N T P^N$ where $\{P^N\}$ is the sequence of orthogonal projections defined in Remark 5.1. Then it is easy to show that $P^N T P^N$ converges in norm to T which implies $\{P^N T P^N\}$ is collectively compact. Thus Theorem 3.2 applies to this type of approximation.

Remark 5.4

From Theorem 2.1 it follows that

$$\mathcal{D}(A^2) = \{ (D\phi, \phi) \in Z \mid \ddot{\phi} \in L_2 \text{ and } D\dot{\phi} = L\phi \}$$

and

$$A^2\phi = (L\dot{\phi}, \ddot{\phi}) \quad \text{for } \phi \in \mathcal{D}(A^2).$$

And it is equivalent to the following generalized boundary value problem:

$$\mathcal{D}(\tilde{A}) = \{ \phi \in L_2 \mid \ddot{\phi} \in L_2 \text{ and } D\dot{\phi} = L\phi \},$$

$$\tilde{A}\phi = \ddot{\phi} \quad \text{for } \phi \in \mathcal{D}(\tilde{A}).$$

Note that $\lambda^2 \in \sigma(\tilde{A})$ if $\lambda \in \sigma(A)$. Thus we may extend the methods in [12], [14] for elliptic boundary value problem to such a problem.

REFERENCES

- [1] P. M. Anselone, Collectively Compact Operator Approximation Theory, Prentice-Hall Englewood Cliff, NJ, 1971.
- [2] H. T. Banks and J. A. Burns, Hereditary Control Problem: Numerical methods based on averaging approximations, SIAM J. Control Optim., 16 (1978), 169-208.
- [3] H. T. Banks, J. A. Burns, and E. M. Cliff, Parameter estimation and identification for systems with delays, SIAM J. Control Optim., 19 (1981), 791-828.
- [4] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, J. Differential Equations, 34 (1979), 496-522.
- [5] H. T. Banks and A. Manitius, Projection series for retarded functional differential equations with applications to optimal control problems, J. Differential Equations, 18 (1975), 296-332.
- [6] J. A. Burns, T. L. Herdman, and H. W. Stech, Linear functional differential equations as semigroup on product spaces, SIAM J. Math. Anal., to appear.
- [7] J. F. G. Francis, The QR transformation: a unitary analogue to the LR transformation, I and II., Computer J., 4 (1961/62), 265-271 and 332-345.

- [8] J. S. Gibson, Linear-quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equations and numerical approximations, SIAM J. Control Optim., to appear.
- [9] K. Ito and T. J. Tarn, Feedback control of linear neutral system, preprint.
- [10] K. Ito and T. J. Tarn, Filtering and stochastic control for linear autonomous neutral systems, preprint.
- [11] F. Kappel, Approximation of neutral functional differential equations in the state space $\mathbb{R}^n \times L^P$, preprint.
- [12] W. Kolata and J. Osborn, Non-selfadjoint spectral approximation and the finite element method, Lecture Notes in Mathematics, 701 (1979), 115-133.
- [13] A. Manitius and R. Triggiani, Function space controllability of linear retarded systems: a derivation from abstract operator conditions, SIAM J. Control Optim., 16 (1978), 599-645.
- [14] W. H. Mills, Jr., The resolvent stability condition for spectra convergence with application to the finite element approximation of noncompact operators, SIAM J. Numer. Anal., 16 (1979), 695-703.
- [15] C. B. Moler and G. W. Stewart, An algorithm for generalized matrix eigenvalue problems, SIAM J. Numer. Anal., 7 (1973), 241-256.

- [16] J. E. Osborn, Spectral approximation for compact operators, Math. Comp. 29 (1975), 712-725.

- [17] R. S. Varga, Functional analysis and approximation theory in numerical analysis, SIAM Publications, 1971.

- [18] R. B. Vinter, On the evolution of state of linear differential delay equations in M^2 : Properties of generator, J. Inst. Math. and Appl. 21 (1978), 13-23.